

Anders Häggman

Department of Mechanical Engineering,
Massachusetts Institute of Technology,
77 Massachusetts Avenue, 3-446,
Cambridge, MA 02139
e-mail: haeggman@mit.edu

Geoff Tsai

Department of Mechanical Engineering,
Massachusetts Institute of Technology,
77 Massachusetts Avenue, 3-446,
Cambridge, MA 02139
e-mail: gtt@mit.edu

Catherine Elsen

LUCID-ULG,
University of Liège,
Chemin des Chevreuils 1, bat. B52,
Liège 4000, Belgium
e-mail: catherine.elsen@ulg.ac.be

Tomonori Honda

Mem. ASME
Department of Mechanical Engineering,
Massachusetts Institute of Technology,
77 Massachusetts Avenue, 3-446,
Cambridge, MA 02139
e-mail: tomonori@mit.edu

Maria C. Yang¹

Fellow ASME
Department of Mechanical Engineering,
Massachusetts Institute of Technology,
77 Massachusetts Avenue, 3-449B,
Cambridge, MA 02139
e-mail: mcyang@mit.edu

Connections Between the Design Tool, Design Attributes, and User Preferences in Early Stage Design

Gathering user feedback on provisional design concepts early in the design process has the potential to reduce time-to-market and create more satisfying products. Among the parameters that shape user response to a product, this paper investigates how design experts use sketches, physical prototypes, and computer-aided design (CAD) to generate and represent ideas, as well as how these tools are linked to design attributes and multiple measures of design quality. Eighteen expert designers individually addressed a 2 hr design task using only sketches, foam prototypes, or CAD. It was found that prototyped designs were generated more quickly than those created using sketches or CAD. Analysis of 406 crowdsourced responses to the resulting designs showed that those created as prototypes were perceived as more novel, more aesthetically pleasing, and more comfortable to use. It was also found that designs perceived as more novel tended to fare poorly on all other measured qualities. [DOI: 10.1115/1.4030181]

Introduction

The goal of product design and development is to create products that fulfill user needs so that consumers will desire and purchase them. In early stage design, design teams generate several design alternatives, then select among them to determine one to pursue for further development [1]. A user-centered strategy to help teams select a design direction is to elicit feedback from users and other stakeholders on provisional design concepts. The design team may then incorporate this feedback into future iterations of the design. This phenomenon of obtaining feedback on provisional design representations has become even more prevalent through the rise of online crowdfunding sites, such as Kickstarter, that present consumers with preproduction designs in order to attract financial investment. Low-cost, quick prototypes, known as “minimum viable product” designs, have been embraced by entrepreneurs as a means to prevalidate business ideas with potential customers [2].

A myriad of factors can play into a user’s responses to a provisional design, from the design’s functionality to its visual styling to the way in which a design is presented to the user. This study

examines and compares two factors that can influence the way a user evaluates a design.

First, this study considers the tools to create a provisional design during the exploratory, generative stage of the design process. A range of design tools may support the development of preliminary concepts, such as 2D sketches, 3D physical prototypes, and digital models, and may do so at different levels of fidelity—from rough representations to realistic renderings. Such tools have inherent capabilities and limitations, which means the same concept created using different tools can result in different designs and thereby potentially influence the feedback that users provide. For example, a preliminary design with complex curves that may be relatively fast and easy to sketch or shape from a piece of foam may be challenging to model using CAD. Moreover, the choice of design tool is in tension with the resources required to create the design representation. Generally, the higher the fidelity of the representation, the more skill and time required to create it. Higher fidelity representations may also require that the designer make additional decisions about design details in order to achieve the desired level of representation fidelity.

Second, this study examines the attributes of the design itself, which may relate to the design’s functionality, interactions, appearance, and use, among others. Key product attributes are not only what users look for when making a purchase decision, but can characterize what it means to be an innovative product [3]. For example, gas mileage may be the most important attribute to a

¹Corresponding author.

Contributed by the Design Theory and Methodology Committee of ASME for publication in the JOURNAL OF MECHANICAL DESIGN. Manuscript received September 19, 2014; final manuscript received March 22, 2015; published online May 19, 2015. Assoc. Editor: Carolyn Seepersad.

car buyer, while screen size may be an important determinant to someone selecting a mobile phone.

This study investigates the interplay between the tools used by practitioners during preliminary design, a product's attributes, and user evaluations of a design, and aims to uncover significant relationships among these using relative, rather than absolute, comparisons. The following research questions are framed:

- How does the choice of design tool impact the rate of idea generation and the total number of ideas produced?
- What is the relationship between the choice of design tool and how users evaluate a design based on its qualities?
- What is the relationship between a product's attributes and its perceived qualities? Are certain design attributes more, or less, strongly linked to specific product qualities?
- What is the interplay of the tools used to create a preliminary design and the attributes of the resulting designs?

Related Work

There is diverse research across design, marketing, and psychology devoted to determining the product features that users will find desirable, including strategies such as conjoint analysis [4] and user-centered design [5,6]. This literature review will not attempt to contextualize that entire body of work, but instead concentrate on subsets that examine the design tools used to create design concepts, the factors that inform how users perceive a design, and ways that early stage design concepts can be evaluated.

Influence of Design Tools Used on the Process of Designing. A substantial set of literature exists on the role of design tools in the early stages of the design process. This section will focus on free-hand sketching, 3D CAD modeling, and the creation of physical prototypes.

Sketching. Sketching design concepts by hand has been found to be an effective technique for early stage design across domains [7]. Sketches are fast to create, and thus permit efficient problem and solution exploration at different levels of abstraction [8]. Sketching enables unexpected discoveries during the process of design [9], and specifically encourages the creation of “see-transform-see” mechanisms for exploration [10]. Sketching can preserve ambiguity while exploring alternatives for a design [11]. Increased visual ambiguity leaves room for uncertainty that facilitates flexible transformations and interpretations which in turn prevents premature commitment to uncreative solutions [12]. However, Stacey and Eckert caution that it is important to distinguish between desirable early stage design ambiguity and undesirable ambiguity in the way a design is communicated [13]. In contrast to much of the above research, a study of expert designers suggests that sketching is not essential for design [14].

CAD Tools. CAD tools are ubiquitous in engineering and product design, but there are questions about their appropriateness during the earliest stages of design. Ullman et al. [15] found that the use of CAD encouraged a depth rather than a breadth approach for the generation of ideas. In surveys, CAD users have noted that the use of CAD too early on can sometimes lead to premature fixation [16]. In situ observation of CAD in the industrial design workplace showed ways in which designers deviate from standard CAD use in order to complement the use of sketches [17]. Fixson and Marion [18] found that adoption of CAD tools too early in the process seemed to lead to a focus on detailed design at the expense of concept development. In a comparison of novice and expert designers, Veisz et al. [19] noted a wide range of beliefs about when both sketching and CAD should be adopted in the design process.

Physical Prototypes. Previous research on the use of physical prototypes in the early stages of design has investigated the simplicity of prototypes [20], the value of low-fidelity prototypes in

reducing uncertainty [21], and as a point of focus for design in teams [22,23]. Houde and Hill [24] delineated prototypes by the type of information that the designer can learn from them: look-and-feel prototypes approximate appearance, implementation prototypes relate to function, and role prototypes offer insight into how a design fits into a user's life.

Comparisons of Design Tools. A body of literature is concerned with comparing paper-based and digital design tools, while physical prototyping is less studied. A study of the use of paper-based tools to prepare for designs that would eventually become digital observed differences in the amount of time spent, though the quality was the same [25]. A comparison between digital drawing and traditional sketching found that traditional tools had advantages in the way concepts were explored and conceived [26]. Stones and Cassidy [27] found that paper-based sketches were better than digital in facilitating idea reinterpretation. A comparison of digital pen, tablet, and CAD found that choice of tool related to the time spent on the design task [28].

Influence of a Product's Perception on User Assessment. The field of industrial design has long considered the instrumental role of a design's appearance in a user's perception of a product—considering not just a product's styling but the broader visual intent of the design. Bloch includes psychological and behavioral components in describing how visual design impacts what consumers want [29]. Crilly et al. [30] formulated a framework for consumer response to the visual that divides that interaction into one between producer and consumer. Strategies have been explored for mapping a product's semantics into a user's perceptual space [31]. There can be variance between what designers intend and what users perceive when viewing a product [32]. Surveys of user perceptions indicated a relationship between the desire to own a product and how a product was perceived [33].

A design tool can influence two key aspects of user perception: representation mode and fidelity. Representation mode refers to the way that concepts are presented, such as photographs, sketches, or renderings. Fidelity refers to the level of detail or realism of the presented designs.

Mode of Representation. Artacho-Ramirez et al. [34] found that as a representation mode became more sophisticated, the differences among how people perceived products decreased. Reid et al. [35] presented a design as computer sketches, computer renderings, and silhouettes and noted variations in consistency of user assessments. Söderman [36] compared sketches, virtual reality, and an actual model, and found that the level of realism played a role in participants' certainty about attributes. Tovaes et al. [37,38] developed a strategy that captures user preferences based on their immediate experiences with a product, as with a virtual model.

Fidelity of Representation. Macomber and Yang [39] focused on levels of fidelity in sketching and CAD and found that realistic hand drawings ranked higher than lower-fidelity sketches or CAD models. Hannah et al. [40] presented low- and high-fidelity sketches, digital models, and prototypes and found that respondents were more confident in their conclusions when viewing high fidelity prototypes. Viswanathan and Linsey [41] found prototypes that required a higher “sunk cost” to create were associated with reduced generation of novelty and variety of ideas. In user interface design, Sauer and Sonderegger [42] found that fidelity can influence estimation of task completion time. Acuna and Sosa [43] compared prototypes created with and without first sketching, and found that originality was marginally higher when participants sketched before creating prototypes.

Assessment of Design Concepts. A continuing area of research is the evaluation of early stage design concepts. Kudrowitz and Wallace [44] offer a comprehensive discussion of metrics for

concept evaluation. Most strategies evaluate designs on an absolute basis, rather than relative. Evaluation is often conducted through objective measurement of physical or process characteristics, or measurement of quality by raters, individually or by panel, expert, or novice. Crowdsourced ratings of creativity correlated with novelty but not with idea usefulness. Clarity in design representation was linked to higher ratings of creativity. Sylcott et al. [45] propose a “metaconjoint” approach that elicits preference information on both form and function, and uses functional magnetic resonance imaging, or functional MRI (fMRI), data to measure responses. Respondents weighed function more heavily than form of the design using both the metaconjoint and fMRI approaches.

What Is the Gap? Research has shown that the way a design is presented—including both the mode and fidelity of representation—can influence how users evaluate a design. At the same time, the design process demands that appropriate design tools be used to create preliminary designs for evaluation. Design tools should allow for design exploration, as well as efficient use of resources. This study examines the complex interplay between design tools and user assessments, as well as links with product attributes. This study further considers these relationships in a relative way, rather than assuming that an individual design concept can be assessed on an absolute basis. Making relative comparisons permits a broader view of the relative importance of each of the factors being studied.

Methods

Overview. Eighteen experienced engineers and designers (“designers”) were asked to generate concepts using one design tool, “sketching,” “prototyping” with blue foam (as is common practice in industrial design), or “CAD,” to address a design task. The resulting designs were then presented in an online survey to evaluate them on product qualities such as novelty, usefulness, and appearance. In parallel, the resulting designs were assessed to determine a set of product attributes that could be used to describe the space of the resulting designs. Six design experts later assessed all resulting designs on these attributes that could be used to describe the space of the resulting designs. These experts later assessed all resulting designs on these attributes.

Expert Design Participants. Designers were recruited via invitations to design firms in Boston and Belgium, to design-related e-mail lists, and to design graduate students at MIT. Designers ranged from 25 to 50 years old, and had 2 to 25 years of design-related work experience. Based on their expertise, seven participants were assigned to the sketching group, six to prototyping, and five to CAD. Participants were compensated \$20 for involvement in the study, with the possibility of an additional \$75 if their design was deemed the “best” in their respective group. The purpose of the additional \$75 was to provide a real-world incentive to create the best possible design.

The design experiment itself was divided into three sections, with interviews before and after each to collect data and to give participants a short break. Designers were free to leave at any point during the experiment. Sketch and prototype activity was videotaped, while CAD was logged using video screen capture.

Before conducting the experiment, three pilot participants tested the experimental protocol. For the pilot, designers were given 3 × 60 min to create concepts. Including introduction, informed consent, and interviews, the total time spent was 4 hr per participant which all pilot participants indicated was too long. Based on this, the experiment time was shortened to 3 × 40 min sessions.

Description of the Design Task. Participants were asked to create at least one design for a remote control for a living room entertainment center. Designers could submit a maximum of three concepts for the competition and were not given any instruction on the type and fidelity of representations that they should

produce. CAD and prototyping participants were also told that they would have an opportunity to explain their ideas to the researchers; the foam or computer models they produced would not have to be self-explanatory.

The remote control was chosen for its familiarity, as well as its relatively low complexity, making it a suitable product for a short design task. The target user group for the remote control was a middle-class family of four (two adults, one teenager, and one small child) who would use the entertainment center 2 hr a day. This entertainment center could include a television, DVD player, digital video recorder (DVR), streaming console, game console, computer, or any other device they felt appropriate:

- Sketch participants were provided letter-sized (for U.S. participants) or A4-sized (for Belgian participants) blank paper and five pencils (2H, 2B, 4B, 6B, 8B), four fineliner markers (0.1 mm, 0.3 mm, 0.5 mm, 0.7 mm), two markers (1.0 mm, 2.0 mm), one chisel tip marker (10.0 mm), a pencil sharpener and eraser.
- Prototype participants were provided as many precut blue foam blocks as they wanted (ranging from 20 cm × 20 cm to 100 cm × 150 cm, with thicknesses from 3 cm to 10 cm), shaping tools (four hand held rasps of varying coarseness), sandpaper (P50, P100, P150, P220), 45 cm long metal ruler, toothpicks (to join foam pieces), glue, a tabletop hot wire cutter (maximum cutting height of 12 cm), and a chisel tip marker. The marker could only be used for marking cut lines on the foam, not for sketching or idea generation purposes.
- CAD participants were provided a desktop computer preloaded with Solidworks modeling software.

Processing Data: Redrawing Designs. At the end of each experiment, sketches were digitally scanned, screenshots were made of CAD models, and photographs were taken of foam models for a total of 83 designs. A standard remote control was also added to the dataset to serve as a baseline reference. The standard remote was the “best-seller” at the time when searching for “remote control” on Amazon.com (Fig. 1).

As has been noted earlier, previous studies have observed that the mode of presentation can influence user perception. Since the

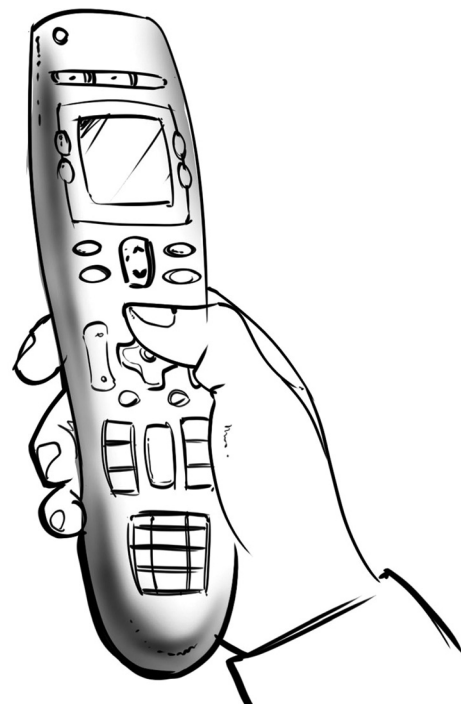


Fig. 1 Sketch of the baseline reference remote control

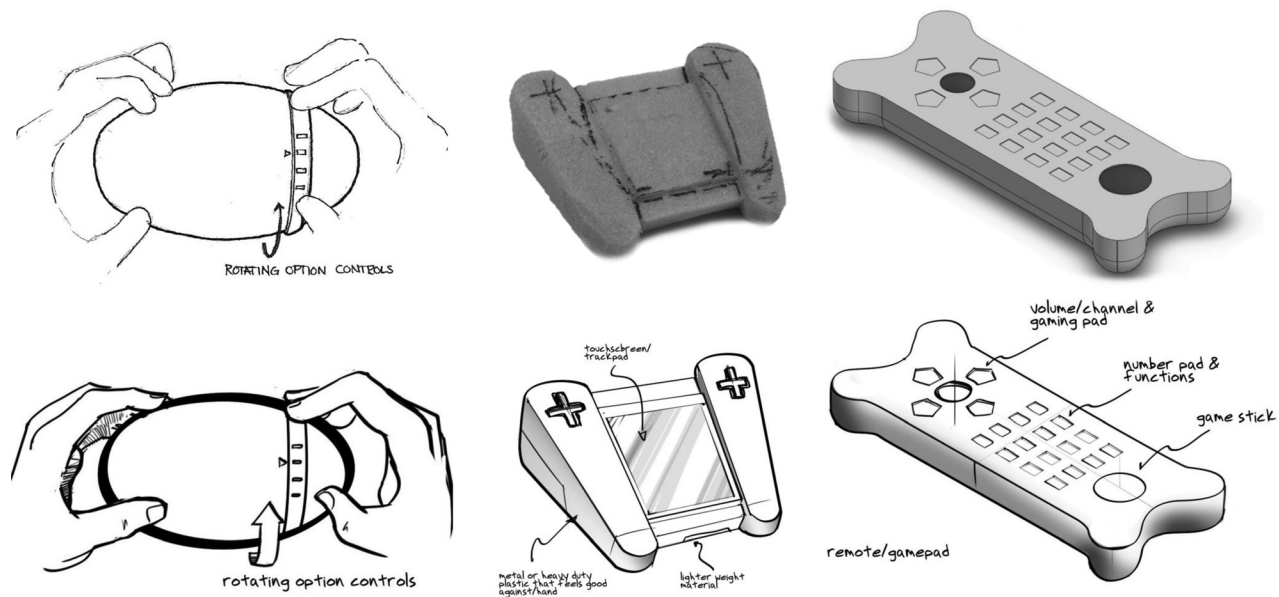


Fig. 2 An original sketch, foam prototype, and CAD model matched with their respective recreated sketch

focus of this study was to compare effects of the design tool in question on the types of concepts generated, all of the ideas created by participants were redrawn as 2D sketches by a professional industrial designer to exclude the effect of the mode of presentation on how an idea was perceived and evaluated. Explanatory annotations based on the interviews with the designers were also added to the redrawn sketches of the foam and computer models in order to make the information content consistent across all three methods—the sketched ideas already included annotations explaining their functionality—and to make the functional principles of the designs understandable to someone seeing them without any further explanation.

The top row of Fig. 2 shows an original sketch, foam prototype, and CAD model for remote controls created by different designers. The bottom row shows the industrial designer's recreation of each.

User Preference Survey

Overview. The resketched concepts were assembled into a survey using Qualtrics (online survey software) and distributed through Amazon Mechanical Turk (an online service for anonymous workers to complete tasks). Mechanical Turk is widely used for social science research and offers a more diverse sample of respondents than a typical college campus sample [46,47]. 506 respondents completed the survey, and after responses from the survey were checked to ensure they were legitimate using quality control questions, 406 responses were accepted.

Survey Design. Ideally, respondents would rank all 83 concepts generated by the designers, but ranking these many concepts would be time consuming and a significant cognitive burden for the respondent. Instead, respondents were presented with a randomly selected subset of the concepts in randomly generated pairs to allow for relative comparisons. Participants were able to respond with their level of preference for Concept A or Concept B using a 5-point scale from “strong preference for A” to “no preference either way, neutral” to “strong preference for B”.

Initially, reviewers were presented with six pairs of images, but based on reviewer feedback on the length of the survey, the number was increased to eight pairs after the first 204 responses were collected. Because the images were randomly chosen, each concept was rated between 58 and 78 times. At the end of the survey, respondents were asked basic demographic information and about

their design-related experience. The survey was designed to take about 15 min to complete.

Each pair of concepts was shown on a single page, with the following questions in random order presented below them:

“Please indicate which of the two concepts you think...”

- looks more useful
- looks more original/creative/novel
- looks more comfortable to use
- you would be more likely to buy (assuming they are similarly priced)
- looks aesthetically more pleasing (looks better)
- is presented more clearly (you understand how the device is meant to work)
- is a better idea (try to give an overall rating, all things considered)

There was also an eighth quality-control question “please click on the strong preference for B option for this question,” the placement of which was random for every pair of images. This is discussed further in the later section on survey quality control.

These rating criteria were chosen based on measures by Garvin's [48] eight dimensions of product quality: performance, features, reliability, conformance to existing product standards, durability, serviceability, aesthetics, and perceived quality. In formulating attributes for this survey, an important consideration was whether a respondent could reasonably make judgments about an attribute based on a line drawing viewed on a computer screen. It was determined that reliability, conformance, durability and serviceability would be difficult to assess in that way. Additionally, these four dimensions were not core to the research questions of this study. The study then focused on performance, features, and aesthetics, with performance expressed as “usefulness” and “comfort during use.”

Survey Quality Control. One of the challenges of collecting anonymous human subjects' data is being confident that the data are legitimate. To accomplish this, only respondents with a 99% approval history on Mechanical Turk were permitted to take the survey. The survey itself also included several questions to ensure high-quality responses. First, at the beginning of the survey, participants were given information about the computer requirements for the survey, and about the design task at hand. On the following pages, they were asked three, simple multiple-choice questions about those requirements. Second, while viewing each pair of

design concepts, one of the questions asked participants to click on the “strong preference for B” option for this question. This question was used to flag users who mindlessly clicked random options, without reading the actual questions. Third, twice during the survey—after a participant had finished rating a pair of images—a required free-response area asked the participant to describe the two concepts previously shown. This question was used to ensure that participants had purposefully considered the images. The time it took for respondents to answer each individual question was also recorded to determine if the respondent had carefully considered the question, or was merely “clicking through” to the next page. All of these methods were used together to determine acceptable responses.

Design Attributes. To establish a set of attributes for the remote control designs, four of the authors independently examined the entire set of designs for common attributes. For example, several designs might include touchscreens, or others buttons. Some designs might require interaction with hands, while others might use only one’s eyes.

Each of the four authors’ sets of attributes was carefully compared, and merged into three categories of attributes: *Form Factor*, *Input*, and *Interaction*. *Form Factor* describes the type of object the design resembles visually. *Input* describes the type of buttons or sensors used in the design—the physical hardware—that allows the user to transmit information to the remote. *Interaction* describes the “primary” type of human interaction required to use the remote, such as “hands.” For example, for a standard remote control (form factor: standard), the input is typically through buttons, while the interaction is with the hands. One could also imagine a remote control shaped like a baseball cap (form factor: novelty/other) that controls a television through brainwaves (input: novelty/other; interaction: novelty/other).

With this set of attributes, a survey was administered to six expert design reviewers twice, with several months in between surveys. Participants in this group had several years experience in design practice, design research or both. In the survey, participants were shown each design concept, and asked to mark the most appropriate attributes and values from a list.

In the first step of attribute analysis, data from the expert surveys were averaged, and concepts were assigned an attribute score based on the level of agreement between experts. For example, a design concept could be 100% interaction with hands, or 0%, or any percentage in-between.

Inter-rater reliability was used to test consistency in mapping each concept sketch into attribute space. Fleiss’ Kappa was chosen as the inter-rater reliability metric because it allows more than two raters [49]. Using Landis and Koch’s criteria [50], it was observed that there was substantial attribute inconsistency among the raters. To address this inter-rater discrepancy, related attributes that were difficult to distinguish were combined. For example, “standard remote” and “game controller” in the “form factor” category. Table 1 provides a complete list of attributes in each of their possible categories.

PCA. The second step of attribute analysis involves Spearman correlation analysis and PCA to determine the amount of coupling and assess the number of distinct attributes. PCA showed that

there was one redundant variable, which makes some sense because sketch, prototype, and CAD are linearly dependent variables. Additionally, there are at least two more dimensions that are most likely redundant. These high correlations and redundancies indicate caution in fitting any kind of model.

Concept Selection. To gain confidence about the mapping between attributes and concept selection, concept selection needs to be evaluated to see if it has a coherent pattern. For example, if concept A is preferred over concept B by half the population, and concept B is preferred over concept A by the other half, it does not make sense to find key attributes to explain why concept A is preferred over concept B. Note that in this example, the heterogeneity of the population must be examined and the population that captures these divided preferences must be segmented. To accomplish this, three different analyses were performed.

Pairwise Consistency. A consistency check focuses on how consistent a population is on comparing pairs of concepts. The main purpose of this consistency check is to see if segmentation of the population is necessary. If concept A is considered better than concept B by half of the population and vice versa, then the population is heterogeneous and needs to be separated into two homogeneous subsets: one that prefers A over B and another population that prefers B over A. The first consistency check was to determine consistency at the pairwise level. Consistency was defined as a percentage of

$$\frac{\sum_{\text{all pairwise comparison}} \max(\text{count}(a > b), \text{count}(b > a))}{\text{count}(\text{all pairwise comparisons with multiple reviewers})} \quad (1)$$

The consistency metrics were mostly above 85%, which suggests random variation within a single homogenous population, rather than a few distinct heterogeneous populations with drastically different preferences.

Ranking-Based Consistency Check. Discrete Choice Model and other utility and preference models were used to map the attribute space into utility or preference values. The goal was to find a utility-based ranking that explained the concept selection for each of the concept qualities (usefulness, creativity, and so forth).

A Colley matrix based ranking, used for college football rankings and gaining use in academic research, was implemented. It assumes the sample size for comparison is limited, similar to football teams who compete in just 12–13 games per season rather than against all other teams in the pool [51]. The number of results per survey had more variability, as if some teams played 6 games per season, while others played 15 games.

Ranking was also directly optimized. This optimization over ranking became a combinatorial NP-hard optimization problem that was solved numerically using local optimization combined with 100 random, initial guesses.

Discrete Choice Model. The mapping from attributes to a utility value, which will determine the likelihood for concept A to be chosen over concept B, is derived using a discrete choice model. One of the main difficulties associated with this analysis is that the attributes seemed to be highly correlated. Additionally, the goal is to determine the most important attributes rather than focus on model accuracy. Given these restrictions, the following techniques were applied:

- (1) Stepwise feature (attribute) selection to remove unnecessary, correlated variables that contribute minimally to the model until the model exhibits a significant decrease in accuracy.
- (2) At each step, L1 and L2 regularization terms were utilized to reduce the complexity of the model and force the contributions from many of the attributes in the discrete choice

Table 1 Attributes organized by attribute category

Form factor	Input	Interaction
Standard remote control	Buttons	Hands
Smartphone/tablet	Joystick	Body
Game controller	Scroll wheel	Eyes
Mouse	Touchpad/touchscreen	Novelty/other
Novelty/other	Gestural	
	Novelty/other	

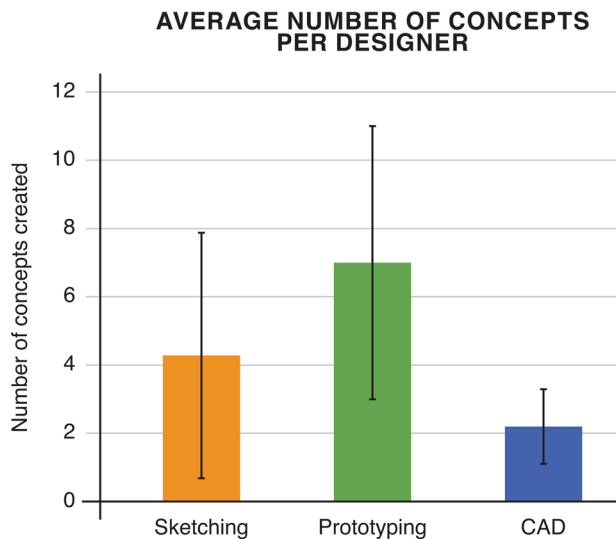


Fig. 3 Average number of concepts per designer, error bars indicate ± 1 standard error

model to be smaller. This aids the stepwise process by revealing which variables are important. L1 and L2 regularizations have been treated as parameters and explored to balance model accuracy with regularized terms. Overfitting was less of a concern given that the number of attributes is comparably small and correlation actually makes the number of independent variables in the principal component space even smaller.

Results and Discussion

Quantity and Time. Of the 83 designs created by the designers, 30 were sketches, 42 foam prototypes, and 11 CAD models. The average number of concepts per designer is shown in Fig. 3. Because of its speed as a design tool, it was expected that sketching would allow designers to generate more ideas in the time allotted than the other two design tools, but instead prototyping led to the largest number of concepts created. Two possible reasons: (1) participants who sketched tended to use less of the allotted 2 hr of time (see Fig. 5) and (2) it was observed that the sketches tended to be polished “communication” type sketches intended to tell a story to an audience, rather than less finished “thinking” sketches meant to enable the designer to reflect and re-interpret. For more

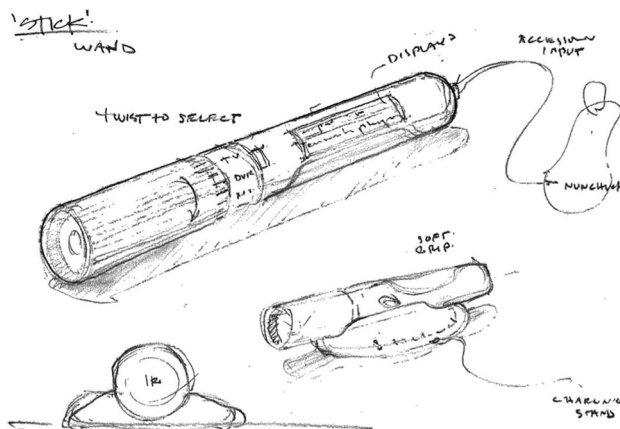


Fig. 4 Example sketch including multiple views and annotations

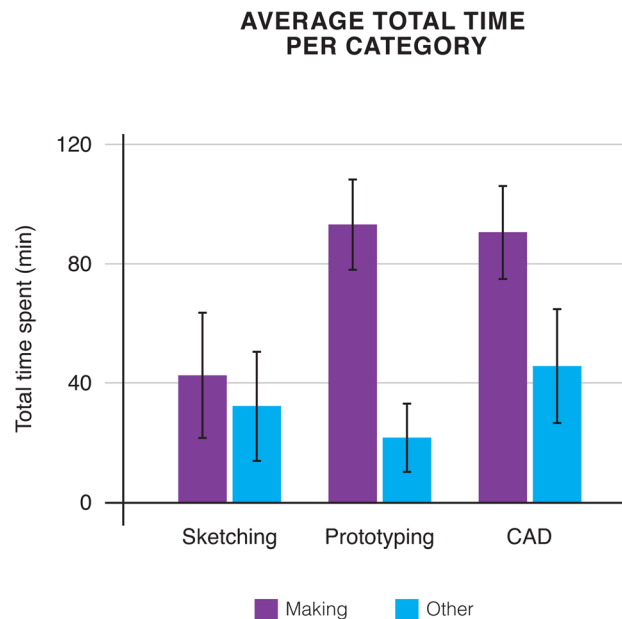


Fig. 5 Average total times spent using each design tool, error bars indicate ± 1 standard error. Making includes time spent actively using specified tool.

explanation concerning differences between thinking, communication, or “talking” sketches, refer to Refs. [52] or [53]. An example of such a communication sketch from the experiment is shown in Fig. 4. It includes different perspectives, annotations, and other details, which presumably means that it took longer to create than a quick thinking type sketch would.

The average total time and time spent per design concept are shown in Figs. 5 and 6. Analyzing video recordings and screen captures of the participants, time spent actively engaged in design (sketching, working with foam, manipulating the CAD model) is labeled “making.” Time spent thinking or evaluating the designs is labeled “other.” CAD clearly required the most time to create a design while prototyping appeared to involve more “active”

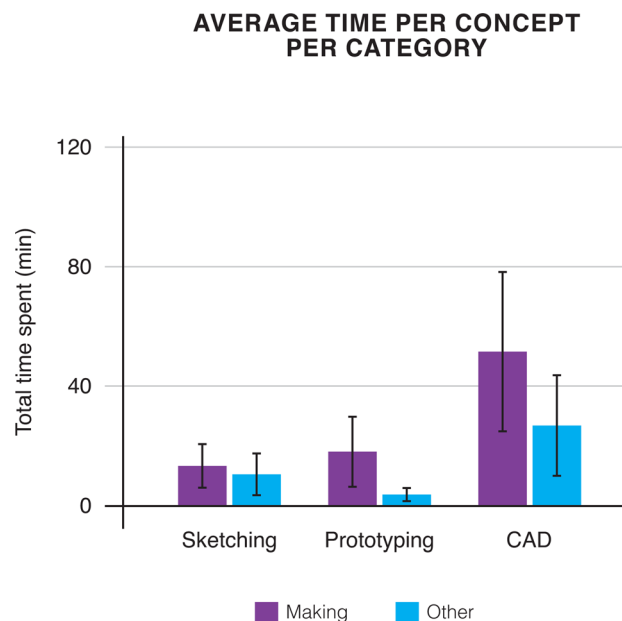


Fig. 6 Average time spent per concept using each design tool, error bars indicate ± 1 standard error. Making includes time spent actively using specified tool.

Table 2 Spearman correlations between attributes and each other, and with design tools. Note that the table is symmetric. Correlations are in **Bold** and *p*-values are in (). *P*-values less than 0.05 have a light gray background.

		Interaction		Form				Input					Design tool		
		Hands and eyes	Body and other	Standard and game	Phone/ tablet	Mouse	Other	Buttons and touch	Joystick	Scroll wheel	Gestural	Other	CAD	Sketch	Proto
Interaction	Hands and eyes		−0.553 (0.000)	+0.346 (0.001)	+0.182 (0.099)	+0.168 (0.128)	−0.411 (0.000)	+0.538 (0.000)	+0.102 (0.357)	+0.173 (0.116)	−0.056 (0.615)	−0.497 (0.000)	+0.112 (0.310)	+0.033 (0.769)	−0.097 (0.379)
	Body and other	−0.553 (0.000)		−0.370 (0.001)	−0.278 (0.011)	−0.082 (0.458)	+0.401 (0.000)	−0.441 (0.000)	−0.165 (0.134)	−0.038 (0.730)	+0.400 (0.000)	+0.570 (0.000)	+0.015 (0.891)	−0.197 (0.073)	+0.152 (0.168)
Form	Standard and game	+0.346 (0.001)	−0.370 (0.001)		−0.043 (0.700)	+0.076 (0.494)	−0.712 (0.000)	+0.607 (0.000)	+0.365 (0.001)	+0.010 (0.926)	−0.191 (0.083)	−0.582 (0.000)	+0.187 (0.088)	+0.099 (0.371)	−0.187 (0.089)
	Phone	+0.182 (0.099)	−0.278 (0.011)	−0.043 (0.700)		+0.043 (0.699)	−0.363 (0.001)	+0.311 (0.004)	−0.091 (0.413)	−0.131 (0.236)	−0.208 (0.058)	−0.385 (0.000)	+0.044 (0.690)	−0.222 (0.042)	+0.170 (0.123)
	Mouse	+0.168 (0.128)	−0.082 (0.458)	+0.076 (0.494)	+0.043 (0.699)		−0.333 (0.002)	+0.230 (0.036)	−0.049 (0.660)	−0.008 (0.946)	+0.275 (0.011)	−0.025 (0.824)	−0.132 (0.231)	−0.167 (0.128)	+0.246 (0.024)
	Other	−0.411 (0.000)	+0.401 (0.000)	−0.712 (0.000)	−0.363 (0.001)	−0.333 (0.002)		−0.701 (0.000)	−0.189 (0.085)	+0.072 (0.514)	+0.087 (0.430)	+0.610 (0.000)	−0.348 (0.001)	+0.108 (0.329)	+0.103 (0.352)
	Buttons and touch	+0.538 (0.000)	−0.441 (0.000)	+0.607 (0.000)	+0.311 (0.004)	+0.230 (0.036)	−0.701 (0.000)		+0.005 (0.967)	+0.053 (0.634)	−0.006 (0.956)	−0.804 (0.000)	+0.255 (0.019)	+0.001 (0.991)	−0.157 (0.155)
Input	Joystick	+0.102 (0.357)	−0.165 (0.134)	+0.365 (0.001)	−0.091 (0.413)	−0.049 (0.660)	−0.189 (0.085)	+0.005 (0.967)		+0.099 (0.373)	−0.056 (0.612)	−0.212 (0.053)	−0.082 (0.458)	+0.197 (0.073)	−0.143 (0.194)
	Scroll wheel	+0.173 (0.116)	−0.038 (0.730)	+0.010 (0.926)	−0.131 (0.236)	−0.008 (0.946)	+0.072 (0.514)	+0.053 (0.634)	+0.099 (0.373)		−0.045 (0.685)	−0.018 (0.869)	+0.212 (0.053)	+0.015 (0.890)	−0.175 (0.111)
	Gestural	−0.056 (0.615)	+0.400 (0.000)	−0.191 (0.083)	−0.208 (0.058)	+0.275 (0.011)	+0.087 (0.430)	−0.006 (0.956)	−0.056 (0.612)	−0.045 (0.685)		+0.126 (0.253)	−0.056 (0.611)	−0.014 (0.902)	+0.034 (0.762)
	Other	−0.497 (0.000)	+0.570 (0.000)	−0.582 (0.000)	−0.385 (0.000)	−0.025 (0.824)	+0.610 (0.000)	−0.804 (0.000)	−0.212 (0.053)	−0.018 (0.869)	+0.126 (0.253)		−0.135 (0.223)	−0.097 (0.378)	+0.159 (0.148)
	Buttons and touch	+0.538 (0.000)	−0.441 (0.000)	+0.607 (0.000)	+0.311 (0.004)	+0.230 (0.036)	−0.701 (0.000)		+0.005 (0.967)	+0.053 (0.634)	−0.006 (0.956)	−0.804 (0.000)	+0.255 (0.019)	+0.001 (0.991)	−0.157 (0.155)
Design tool	CAD	+0.112 (0.310)	+0.015 (0.891)	+0.187 (0.088)	+0.044 (0.690)	−0.132 (0.231)	−0.348 (0.001)	+0.255 (0.019)	−0.082 (0.458)	+0.212 (0.053)	−0.056 (0.611)	−0.135 (0.223)		−0.272 (0.012)	−0.386 (0.000)
	Sketch	+0.033 (0.769)	−0.197 (0.073)	+0.099 (0.371)	−0.222 (0.042)	−0.167 (0.128)	+0.108 (0.329)	+0.001 (0.991)	+0.197 (0.073)	+0.015 (0.890)	−0.014 (0.902)	−0.097 (0.378)	−0.272 (0.012)		−0.760 (0.000)
	Proto	−0.097 (0.379)	+0.152 (0.168)	−0.187 (0.089)	+0.170 (0.123)	+0.246 (0.024)	+0.103 (0.352)	−0.157 (0.155)	−0.143 (0.194)	−0.175 (0.111)	+0.034 (0.762)	+0.159 (0.148)	−0.386 (0.000)	−0.760 (0.000)	

Table 3 Rank accuracy summary

	Useful	Creative	Comfortable	Buy	Looks	Clarity	Better idea
Colley rank weighted accuracy	0.779	0.716	0.760	0.756	0.731	0.764	0.764
Optimized rank weighted accuracy	0.817	0.751	0.791	0.793	0.776	0.803	0.795

engagement with the material and tools as a percentage of overall time.

Relationship Between Product Qualities and Design Tools. Table 2 shows Spearman correlations between attributes themselves and with design tools to help evaluate consistency within a design concept. Correlations are in bold text, while p -values are in parentheses; additionally, those with $p < 0.05$ have a light gray background. Note that the matrix is symmetric between the attributes, though the full set of correlations is shown for the sake of convenience. For example, forms that took the shape of standard remote and game controller showed a positive correlation with both input from “buttons and touchscreen/touchpad” and with “joystick” (correlation, p -value: +0.607, 0.000, and +0.365, 0.001, respectively). This makes sense; it is expected that standard remotes and game controllers would have these types of controls. Similarly, there was a positive correlation with interactions that involved “body and novelty/other” with “novelty/other” forms (correlation, p -value +0.401, 0.000). Again, this is logical because designs that do not have traditional types of interaction—for example, using body movement or brain waves—would likely be paired with unconventional forms, i.e., not standard remote and game controller, “smartphone/tablet,” or “mouse.” In addition, both novelty/other forms and body and novelty/other interactions are positively correlated with other input, further supporting this notion (correlation, p -value: +0.610, 0.000, and +0.570, 0.000, respectively).

Links between attributes and design tools? It was found that sketchers did not generally create smartphone or tabletlike forms (correlation, p -value −0.222, 0.042), and that mouse forms tended to be created using foam prototypes. Designs created using CAD tended to include buttons and touchpads as input, which makes sense because CAD tools are well suited to modeling such features. A particularly interesting finding is that CAD designs tended not to be used to create forms categorized as novelty/other. Other research cited in this paper finds that adopting CAD too

early in the design process causes designers to limit their concept exploration prematurely. This study’s finding suggests that early stage CAD is linked with designs that are not novel as well, a result that could possibly be linked with premature fixation.

Relationship Between Representation and Design Qualities: Top Designs. Another way to examine how the tool used influences the design is to determine concepts rank the highest on a particular design quality. This approach of looking at the highest ranked designs makes sense given the context of a design process where multiple designs are generated but only the best ideas survive to become further developed. To accomplish this, Colley ranking and optimized rankings were applied to the user comparison data. Table 3 shows the weighted accuracies of the Colley and optimized rankings.

Colley ranking was developed as a method of ranking for the U.S. College Football Bowl Championship Series system. One of the difficulties of ranking college football teams is the unbalanced schedule and small sample size. An unbalanced schedule means that some teams play a “tough” schedule (playing mostly against better teams), while some teams play a “soft” schedule (against weaker teams). A team that plays a soft schedule might have fewer losses, but if they were switched to a tough schedule they might not win as often. This scenario is similar to the pairwise comparisons from the survey.

Every survey comparison is treated like the outcome of a football game and applied the Colley ranking algorithm [51]. Then the probability is computed that a given team will win against an opponent, considering their opponent’s strength. As the sample size increases, the schedule becomes more balanced. The following formula is used to compute the final ranking accuracy:

$$\text{ranking accuracy} = \frac{(\text{count of higher} - \text{ranked concept winning})}{(\text{total number of comparison})} \quad (2)$$

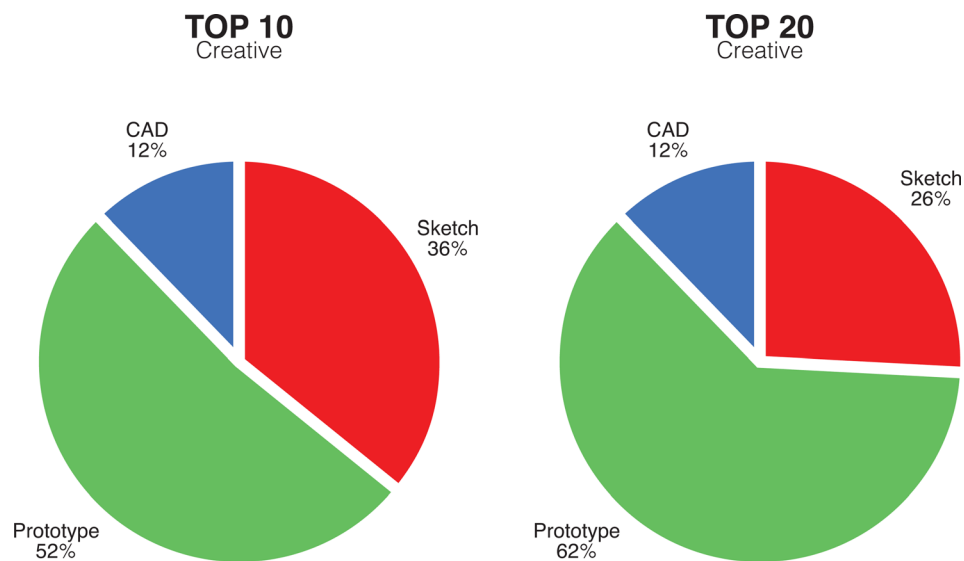


Fig. 7 The most creative designs, normalized by the number of participants

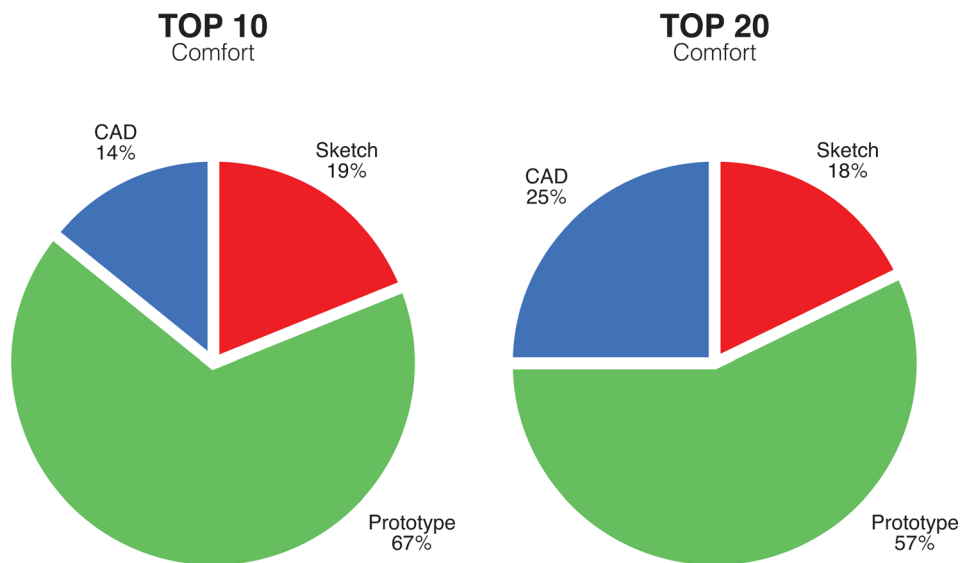


Fig. 8 The most comfortable looking designs, normalized by the number of participants

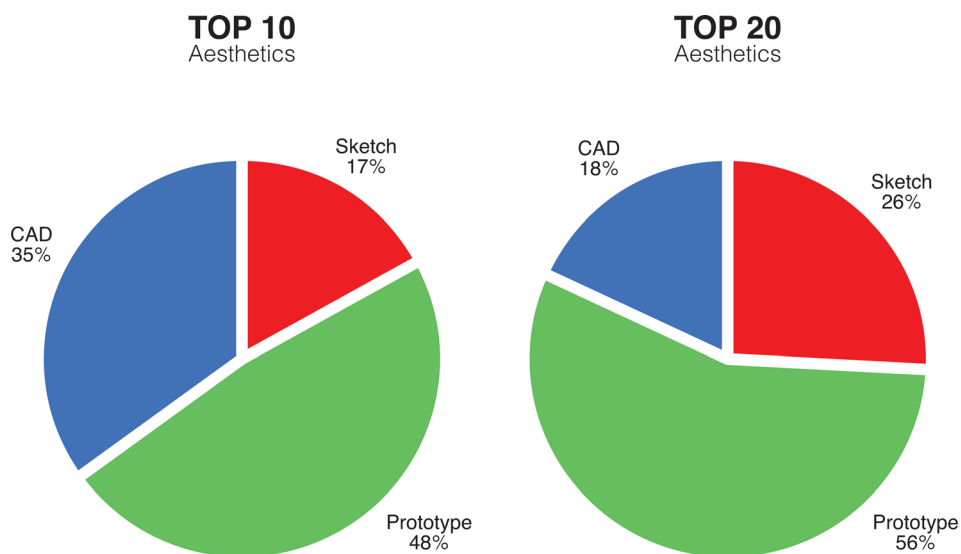


Fig. 9 The most aesthetically pleasing designs, normalized by the number of participants

If there is no inherent difference between concepts, this accuracy should be around 50%.

For optimized ranking, a brute force heuristic optimization technique is applied to the Colley ranking to improve the final ranking accuracy, optimizing the ranking of concepts such that ranking accuracy is maximized. This provides an upper bound on the discrete choice modeling accuracy given the data set. This is because the discrete choice model maps from attribute space into utility and determines the likelihood by comparing utility values of two concepts. The optimized ranking actually reflects the ideal ranking on utility space that the discrete choice model should map into.

The charts (Figs. 8 and 9) show the top ten and twenty ranked designs as shown in the optimized rankings. As a point of comparison, out of the top ten ranked concepts between two and eight concepts were the same, regardless of whether the Colley or optimized ranking method was used. This overlap was particularly notable for aesthetics, clarity, and selection as the “better” design.

Figure 7 shows the top-ranked creative designs, normalized by the number of participants per type of design tool. Because there were different numbers of participants using sketching,

prototyping, and CAD, the number of ideas in the top ranking for creativity was divided by the number of participants who used that tool. Then, the normalized results were represented as a percentage of the whole—in the top ten and top twenty. The red area indicates sketching, green prototyping, and blue CAD.

Figure 7 shows how designs produced with foam models (prototyping) dominated the top-ranked creative designs. It was expected that sketching would have produced a larger share of design concepts perceived to be most creative because of the facility and speed with which a participant could explore the design space. However, implicit in that expectation is the idea that sketchers would use fast-to-create thinking drawings rather than the slower, more detailed communication drawings they actually produced. At the same time, prototypers generally created models with limited details, presumably because of the difficulty of creating intricate details with blue foam. This is not to say the prototypers did not envision detailed designs; their interviews indicated that they had in mind a detailed view of their designs. Because foam is suited to rough, low-fidelity modeling, participants were able to generate many concepts quickly.

Table 4 Variable importance of design attributes and design tools to design quality measures

		Useful	Creative	Comfortable	Buy	Aesthetics	Clarity	Better idea
	Weighted accuracy	0.7077	0.6492	0.6663	0.7113	0.6579	0.6991	0.7095
Interaction	Hands and eyes	0.0	0.0	0.0	+24.9	0.0	+122.4	0.0
	Body and novelty/other	0.0	+49.4	0.0	0.0	-0.4	+131.1	-0.8
Form	Standard remote and game controller	+535.8	0.0	+699.0	0.0	0.0	0.0	0.0
	Smartphone/tablet	0.0	+316.8	-15.1	+1327.0	+3419.5	+1.8	+928.7
	Mouse	0.0	0.0	-18.2	+6.5	0.0	+31.5	0.0
	Novelty/other	-704.5	+715.4	-691.2	-709.5	-682.1	-616.6	-709.0
Input	Buttons and touchpad/touchscreen	+0.1	-709.8	0.0	0.0	+1.2	0.0	+57.3
	Joystick	-64.2	0.0	0.0	0.0	0.0	0.0	0.0
	Scroll wheel	0.0	0.0	0.0	-43.2	+0.0	-22.4	0.0
	Gestural	+11.5	0.0	0.0	+83.9	-0.8	+9.6	0.0
	Novelty/other	0.0	0.0	-4.3	0.0	-27.3	-199.0	0.0
Design tool	Sketching	0.0	0.0	-14.7	0.0	0.0	-24.5	-0.4
	Prototyping	0.0	0.0	0.0	0.0	0.0	-14.5	0.0
	CAD	0.0	0.0	+23.0	+67.3	0.0	-24.0	+0.1

Figures 8 and 9 show the breakdown of tools used to create the top comfortable and “aesthetically” pleasing designs. Again, prototyping dominates the top-ranked designs, suggesting the value of low fidelity representations on these qualities. With respect to “useful,” “likely to buy,” “clarity,” and “better idea,” no design-tool clearly dominated the top designs.

The one design not represented in these charts is the standard remote, which was not generated with a specific tool. Not surprisingly, the standard remote did not rank highly for *creative* but it did rank in the top 20 for *likely to buy*, the top 15 for *better overall design*, and the top 10 for *useful*.

Design Attributes, Design Tools, and Qualities. This section links together all three design variables of interest: design tools, their perceived qualities, and the attributes of the design concepts. Table 4 shows the relative importance of design attributes and design tools with respect to each of the design quality measures calculated using discrete choice modeling. For a given column, each cell can be read relative to each other. Orders of magnitude differences are meaningful, and shading is graduated to reflect this, as in a heat map. Columns should not be compared with each other.

Design Tools and Perceived Design Qualities. Designs created in CAD were perceived as slightly more “comfortable” than those created using sketches. Designs created in CAD were also judged as “more likely to buy” than those created by other tools. Figure 10 shows a design created in CAD that was perceived as more likely to buy as ranked in both the Colley and optimized rankings. This was somewhat unexpected because the physical

form itself is a simple rectangular block. This design was notable in that it was a software app that could be downloaded to a smartphone rather than a dedicated remote control device. A few possible reasons for this result: it could be that part of the purchase appeal was that this particular smartphone was perceived to be an Apple iPhone and therefore deemed to be desirable via its association with the brand rather than because of the intrinsic value of the design itself, or that apps tend to be less expensive than dedicated remotes.

In terms of a design’s perceived clarity, it can be seen that CAD, sketch, and prototype all have negative values. These negative values are due to the regularization used during discrete choice modeling, and would not normally happen if the variables were independent. Because of dependencies between CAD and the attribute Form: Novelty/other, *prototype* with the attribute Form: Mouse, and *sketch* with the attribute Form: Smartphone, negative values reflect mostly second-order effects. Overall, relationships between the design tools and qualities are relatively small in magnitude in comparison to the attributes.

Design Attributes and Perceived Design Qualities. Table 4 shows that designs judged as novel (form: novelty/other) had a tendency to be perceived negatively on all qualities except creativity, which it had a strongly positive association with. Novelty appeared to have a negative link with clarity, suggesting that respondents did not necessarily understand how creative designs functioned. To illustrate, Fig. 11 shows a novelty design (form: Novelty/other) that respondents perceived as original/creative/novel as ranked by both the Colley and optimized rankings. This is a remote that can be controlled by a user’s brain waves.

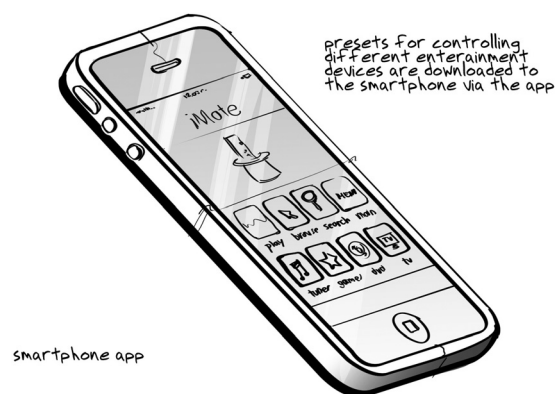


Fig. 10 Example of a design created in CAD that has been resketched

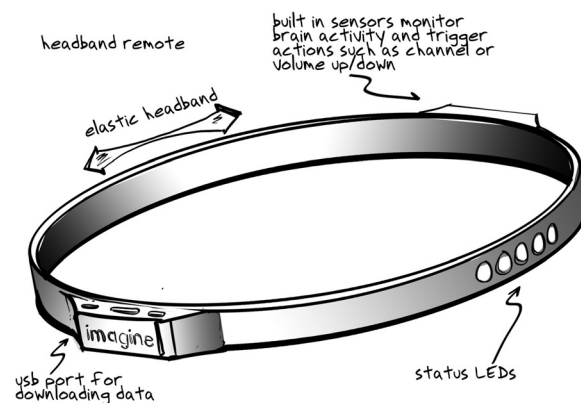


Fig. 11 Example design with novelty/other form and high creative/novel quality

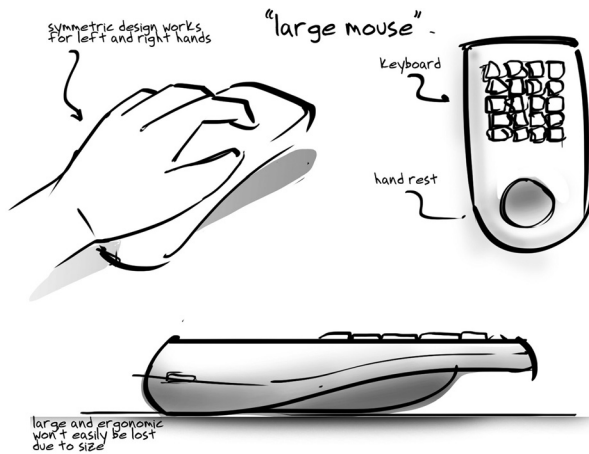


Fig. 12 Example design with buttons or touchpad as input and low creative/novel quality

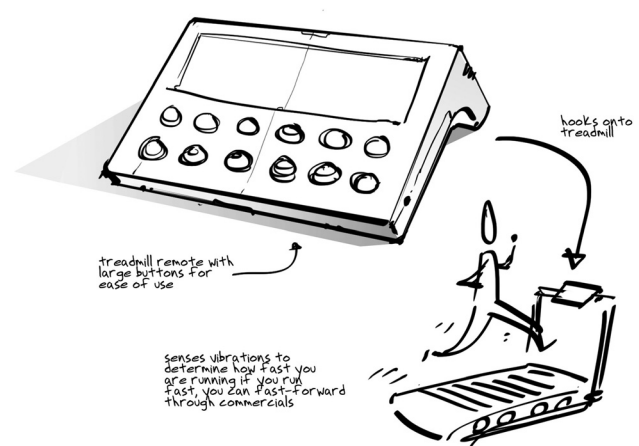


Fig. 14 Example design with body and novelty/other interaction and a high clarity quality



Fig. 13 Example design with a standard remote form with high useful and comfortable qualities

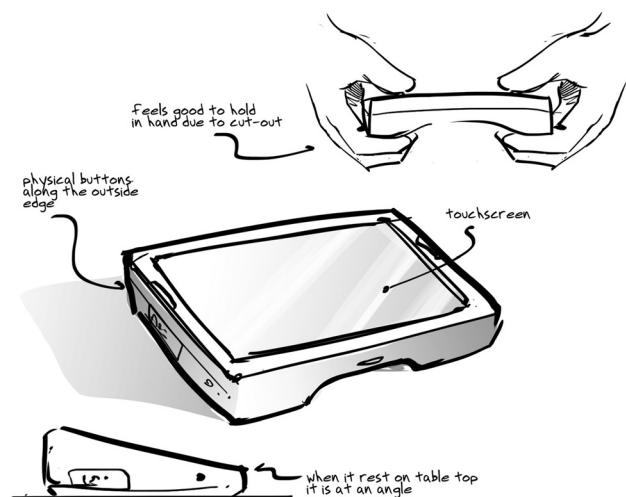


Fig. 15 Example design with a smartphone/tablet form and high aesthetics quality

Respondents felt it was creative, but one could imagine that its operation was ambiguous, or implausible.

In contrast, Fig. 12 shows a design that included buttons (Input: Buttons & Touchpad) and was perceived as not creative. This perception could be because the design is easily recognizable as a mouse or keyboard-style input.

Standard remotes and game controllers (form: Standard remote and game) were evaluated as both useful and comfortable (Fig. 13). This makes sense given that these are forms that respondents are likely familiar with and have been designed specifically for use as remote controls.

Designs that involved interaction with the body itself, rather than hands or the eyes (interaction: body and novelty/other) were perceived as being clear. Figure 14 shows an example in which a remote control is operated by a user on a treadmill.

Finally, aesthetics are considered. Designs that were classified as smartphones or tablets (form: smartphone/tablet) were strongly perceived as aesthetically pleasing (Fig. 15). In fact, smartphones were also perceived positively for originality, for purchase, and overall considered a better idea.

Conclusions

This paper explored the role of the design tool used for early design exploration, product quality and product attributes. Key

findings related to each research question are highlighted and discussed in response to the original research questions:

- How does the choice of design tool impact the rate of idea generation and the total number of ideas produced?

Key finding: Foam prototyping resulted in faster generation of ideas than sketching or CAD.

Working with foam prototypes produced more ideas more quickly than with sketching. While sketching is generally a fast, flexible tool for design representations, in this experiment, participants tended to create detailed communication sketches, which took more time than rougher thinking sketches. In contrast, prototypes tended to create fast, low-fidelity prototypes with little detail. The takeaway is not that a particular tool is better than another, but that the level of fidelity of the tools is a crucial factor in speed and quantity regardless of the tool selected:

- What is the relationship between the choice of design tool and how users evaluate a design based on its qualities?

Key finding: In this study, when looking at the top-rated concepts, foam prototypes are perceived positively on a number of qualities: creativity, comfort, and aesthetics.

Of the top concepts, prototyped designs were perceived as having higher novelty than designs created using sketching or CAD, presumably because these tools limited design space exploration

when compared to rough prototyping using blue foam. This result could also be influenced by the evolving interactions between the designer and prototype as part of a “conversation with materials” [10]. Additionally, designs created with CAD were negatively associated with the generation of novel physical forms. This could be due to the constraining nature of CAD used too early in the design cycle:

- What is the relationship between a product’s attributes and its perceived qualities?

Key finding: A novel form alone is sometimes not sufficient for a well perceived design.

Remote controls with forms resembling standard remotes or game controllers were considered useful and comfortable, while smartphone- and tablet-type forms were considered beautiful, novel, more likely to be purchased, and better overall. A somewhat unexpected finding was that concepts with novel physical embodiments were perceived negatively for all other qualities except creativity. A basic assumption in early stage design is that the generation of creative ideas will lead to more desirable design solutions [54]. However, the present study result suggests that novelty by itself does not necessarily mean that a design will be perceived positively on any other measure. Novelty may be a necessary condition for design success, but it is not a sufficient condition on its own.

- What is the interplay of the tools used to create a preliminary design and the attributes of the resulting designs?

Designs created using CAD tended to include buttons and touchpads as input, which was not surprising, but CAD designs tended *not* to be used to create forms categorized as novelty/other.

Future Work

This study focused on a set of design tools that are widely employed in product and industrial design, and future work should broaden this suite of tools to include others such as rapid prototyping. This study looked at attributes from the point of the user. However, designers and engineers need to be able to relate user perceptions to a design’s underlying functional [55] and engineering characteristics as well. More broadly, this study focused on only one aspect of the process, the design of the product itself. However, the design and development of products is a challenging and complex endeavor that must be integrated within a larger context of system-level design, manufacturing as well as a product’s intended market [56,57] and retail channels [58]. Future work should examine how choice of design tools and representation might influence the greater scope of how a product is marketed, distributed, and sold.

Acknowledgment

The authors are grateful for the generous and expert assistance of Alison Olechowski and Catherine Fox.

This work was supported in part by the SUTD-MIT International Design Centre, the National Science Foundation under Award Nos. CMMI-1130791 and CMMI-1334267, and the Finnish Foundation for Technology Promotion.

References

- [1] Ulrich, K. T., and Eppinger, S. D., 2000, *Product Design and Development*, McGraw-Hill, Inc., New York.
- [2] Ries, E., 2011, *The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses*, Random House, LLC, New York.
- [3] Saunders, M. N., Seepersad, C. C., and Hölttä-Otto, K., 2011, “The Characteristics of Innovative, Mechanical Products,” *ASME J. Mech. Des.*, **133**(2), p. 021009.
- [4] Green, P. E., and Srinivasan, V., 1978, “Conjoint Analysis in Consumer Research: Issues and Outlook,” *J. Consum. Res.*, **5**(2), pp. 103–123.
- [5] Norman, D. A., 1988, *The Psychology of Everyday Things*, Basic Books (AZ).

- [6] Faste, R., 1987, “Perceiving Needs,” SAE Technical Paper No. 871534, Society of Automotive Engineers, New York.
- [7] Eckert, C., Blackwell, A., Stacey, M., Earl, C., and Church, L., 2012, “Sketching Across Design Domains: Roles and Formalities,” *Artif. Intell. Eng. Des. Anal. Manuf.*, **26**(Special Issue 03), pp. 245–266.
- [8] Cross, N., 2000, *Strategies for Product Design*, Wiley, London.
- [9] Visser, W., 2006, *The Cognitive Artifacts of Designing*, L. Erlbaum Associates Inc. Hillsdale.
- [10] Schön, D. A., and Wiggins, G., 1992, “Kinds of Seeing and Their Functions in Designing,” *Des. Stud.*, **13**(2), pp. 135–156.
- [11] Goel, V., 1995, *Sketches of Thought*, MIT, Cambridge.
- [12] Tseng, W. S., and Ball, L. J., 2011, “How Uncertainty Helps Sketch Interpretation in a Design Task,” *Design Creativity 2010*, T. Taura and Y. Nagai, eds., Springer, London, pp. 257–264.
- [13] Stacey, M., and Eckert, C., 2003, “Against Ambiguity,” *Comput. Supported Coop. Work (CSCW)*, **12**(2), pp. 153–183.
- [14] Bilda, Z., Gero, J. S., and Purcell, T., 2006, “To Sketch or Not to Sketch? That Is the Question,” *Des. Stud.*, **27**(5), pp. 587–613.
- [15] Ullman, D. G., Wood, S., and Craig, D., 1990, “The Importance of Drawing in the Mechanical Design Process,” *Comput. Graph.*, **14**(2), pp. 263–274.
- [16] Robertson, B., and Radcliffe, D., 2009, “Impact of CAD Tools on Creative Problem Solving in Engineering Design,” *Comput.-Aided Des.*, **41**(3), pp. 136–146.
- [17] Elsen, C., Darses, F., and Leclercq, P., 2011, “An Anthro-Based Standpoint on Mediating Objects: Evolution and Extension of Industrial Design Practices,” *Design Computing and Cognition'10*, J. Gero, ed., Springer, Amsterdam, pp. 55–74.
- [18] Fixson, S. K., and Marion, T. J., 2012, “Back-loading: A Potential Side Effect of Employing Digital Design Tools in New Product Development,” *J. Prod. Innovation Manage.*, **29**(S1), pp. 140–156.
- [19] Veisz, D., Namouz, E. Z., Joshi, S., and Summers, J. D., 2012, “Computer-Aided Design Versus Sketching: An Exploratory Case Study,” *Artif. Intell. Eng. Des. Anal. Manuf.*, **26**(Special Issue 03), pp. 317–335.
- [20] Yang, M. C., 2005, “A Study of Prototypes, Design Activity, and Design Outcome,” *Des. Stud.*, **26**(6), pp. 649–669.
- [21] Gerber, E., 2009, “Prototyping: Facing Uncertainty Through Small Wins,” International Conference on Engineering Design (ICED’09), Stanford.
- [22] Edelman, J. A., Leifer, L., Banerjee, B., Sonalkar, N., Jung, M., and Lande, M., 2009, “Hidden in Plain Sight: Affordances of Shared Models in Team Based Design,” 17th International Conference on Engineering Design (ICED 09), Vol. 2, Design Theory and Research Methodology, Aug. 24–27, Palo Alto.
- [23] Schrage, M., and Peters, T., 1999, *Serious Play: How the World's Best Companies Simulate to Innovate*, Harvard Business School, Boston.
- [24] Houde, S., and Hill, C., 1997, “What do Prototypes Prototype?,” *Handbook of Human-Computer Interaction*, M. Helander, T. Landauer and P. Prabhu, eds., Elsevier Science, Amsterdam.
- [25] Black, A., 1990, “Visible Planning on Paper and on Screen: The Impact of Working Medium on Decision-Making by Novice Graphic Designers,” *Behav. Inf. Technol.*, **9**(4), pp. 283–296.
- [26] Bilda, Z., and Demirkan, H., 2003, “An Insight on Designers’ Sketching Activities in Traditional Versus Digital Media,” *Des. Stud.*, **24**(1), pp. 27–50.
- [27] Stones, C., and Cassidy, T., 2010, “Seeing and Discovering: How Do Student Designers Reinterpret Sketches and Digital Marks During Graphic Design Ideation?,” *Des. Stud.*, **31**(5), pp. 439–460.
- [28] Vasantha, G. V. A., Chakrabarti, A., Rout, B. K., and Corney, J., 2014, “Influences of Design Tools on the Original and Redesign Processes,” *Int. J. Des. Creativity Innovation*, **2**(1), pp. 20–50.
- [29] Bloch, P. H., 1995, “Seeking the Ideal Form: Product Design and Consumer Response,” *J. Mark.*, **59**(3), pp. 16–29.
- [30] Crilly, N., Moultrie, J., and Clarkson, P. J., 2004, “Seeing Things: Consumer Response to the Visual Domain in Product Design,” *Des. Stud.*, **25**(6), pp. 547–577.
- [31] Petiot, J.-F., and Yannou, B., 2004, “Measuring Consumer Perceptions for a Better Comprehension, Specification and Assessment of Product Semantics,” *Int. J. Ind. Ergon.*, **33**(6), pp. 507–525.
- [32] Ahmed, S., and Boelskifte, P., 2006, “Investigations of Product Design Engineering Students Intentions and a Users Perception of Product Character,” NordDesign Conference 2006, Education, Design Methods and Applications, Reykjavik, Aug. 16–18.
- [33] Perez Mata, M., Ahmed-Kristensen, S., and Yanagisawa, H., 2013, “Perception of Aesthetics in Consumer Products,” International Conference on Engineering Design, Design Society.
- [34] Artacho-Ramirez, M., Diego-Mas, J., and Alcaide-Marzal, J., 2008, “Influence of the Mode of Graphical Representation on the Perception of Product Aesthetic and Emotional Features: An Exploratory Study,” *Int. J. Ind. Ergon.*, **38**(11), pp. 942–952.
- [35] Reid, T. N., MacDonald, E. F., and Du, P., 2013, “Impact of Product Design Representation on Customer Judgment,” *ASME J. Mech. Des.*, **135**(9), p. 091008.
- [36] Söderman, M., 2005, “Virtual Reality in Product Evaluations With Potential Customers: An Exploratory Study Comparing Virtual Reality With Conventional Product Representations,” *J. Eng. Des.*, **16**(3), pp. 311–328.
- [37] Tovaes, N., Boatwright, P., and Cagan, J., 2014, “Experiential Conjoint Analysis: An Experience-Based Method for Eliciting, Capturing, and Modeling Consumer Preference,” *ASME J. Mech. Des.*, **136**(10), p. 101404.
- [38] Tovaes, N., Cagan, J., and Boatwright, P., 2013, “Capturing Consumer Preference Through Experiential Conjoint Analysis,” *ASME Paper No. DETC2013-12549*.

- [39] Macomber, B., and Yang, M. C., 2011, "The Role of Sketch Finish and Style in User Responses to Early Stage Design Concepts," *ASME Paper No. DETC2011-48714*.
- [40] Hannah, R., Joshi, S., and Summers, J. D., 2012, "A User Study of Interpretability of Engineering Design Representations," *J. Eng. Des.*, **23**(6), pp. 443–468.
- [41] Viswanathan, V., and Linsey, J., 2011, "Design Fixation in Physical Modeling: An Investigation on the Role of Sunk Cost," *ASME Paper No. DETC2011-47862*.
- [42] Sauer, J., and Sonderegger, A., 2009, "The Influence of Prototype Fidelity and Aesthetics of Design in Usability Tests: Effects on User Behaviour, Subjective Evaluation and Emotion," *Appl. Ergon.*, **40**(4), pp. 670–677.
- [43] Acuna, A., and Sosa, R., 2011, "The Complementary Role of Representations in Design Creativity: Sketches and Models," *Design Creativity 2010*, T. Taura, and Y. Nagai, eds., Springer, London, pp. 265–270.
- [44] Kudrowitz, B. M., and Wallace, D., 2012, "Assessing the Quality of Ideas From Prolific, Early-Stage Product Ideation," *J. Eng. Des.*, **24**(2), pp. 120–139.
- [45] Sylcott, B., Cagan, J., and Tabibnia, G., 2013, "Understanding Consumer Tradeoffs Between Form and Function Through Metaconjoint and Cognitive Neuroscience Analyses," *ASME J. Mech. Des.*, **135**(10), p. 101002.
- [46] Paolacci, G., Chandler, J., and Ipeirotis, P. G., 2010, "Running Experiments on Amazon Mechanical Turk," *Judgment Decis. Making*, **5**(5), pp. 411–419.
- [47] Buhrmester, M., Kwang, T., and Gosling, S. D., 2011, "Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data?," *Perspect. Psychol. Sci.*, **6**(1), pp. 3–5.
- [48] Garvin, D. A., 1984, "What Does 'Product Quality' Really Mean?," *Sloan Manage. Rev.*, **26**(1), pp. 25–43.
- [49] Fleiss, J. L., 1971, "Measuring Nominal Scale Agreement Among Many Raters," *Psychol. Bull.*, **76**(5), pp. 378–382.
- [50] Landis, J. R., and Koch, G. G., 1977, "The Measurement of Observer Agreement for Categorical Data," *Biometrics*, **33**(1), pp. 159–174.
- [51] Colley, W. N., 2002, *Colley's Bias Free College Football Ranking Method: The Colley Matrix Explained*, Princeton University, Princeton.
- [52] Ferguson, E. S., 1992, *Engineering and the Mind's Eye*, MIT, Cambridge.
- [53] Tovey, M., and Richards, C., 2004, "Computer Representation for Concept Design and Maintenance Instruction," *Tools and Methods of Competitive Engineering*, pp. 107–115.
- [54] Kelley, T., and Littman, J., 2001, *The Art of Innovation: Lessons in Creativity From IDEO*, America's Leading Design Firm, Doubleday, New York.
- [55] Hirtz, J., Stone, R. B., McAdams, D. A., Szykman, S., and Wood, K. L., 2002, "A Functional Basis for Engineering Design: Reconciling and Evolving Previous Efforts," *Res. Eng. Des.*, **13**(2), pp. 65–82.
- [56] Michalek, J. J., Ceryan, O., Papalambros, P. Y., and Koren, Y., 2005, "Balancing Marketing and Manufacturing Objectives in Product Line Design," *ASME J. Mech. Des.*, **128**(6), pp. 1196–1204.
- [57] Kumar, D., Chen, W., and Simpson, T. W., 2008, "A Market-Driven Approach to Product Family Design," *Int. J. Prod. Res.*, **47**(1), pp. 71–104.
- [58] Williams, N., Azarm, S., and Kannan, P. K., 2008, "Engineering Product Design Optimization for Retail Channel Acceptance," *ASME J. Mech. Des.*, **130**(6), p. 061402.